# RAW SEQUENCE LISTING
## ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: _10/001, 486 A_

Source: _OIPE_

Date Processed by STIC: _7-8-02_

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:
1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.
PATENTIN 2.1 e-mail help: **patin21help@uspto.gov** or phone 703-306-4119 (R. Wax)
PATENTIN 3.0 e-mail help: **patin3help@uspto.gov** or phone 703-306-4119 (R. Wax)

TO **REDUCE** ERRORED SEQUENCE LISTINGS, PLEASE USE THE **CHECKER VERSION 3.1 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:
**http://www.uspto.gov/web/offices/pac/checker**

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.
Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.
Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. **EFS-Bio (<http://www.uspto.gov/ebc/efs/downloads/documents.htm>** , **EFS Submission User Manual** - ePAVE)
2. **U.S. Postal Service:** U.S. Patent and Trademark Office, Box Sequence, P.O. Box 2327, Arlington, VA 22202
3. **Hand Carry directly to:**
   U.S. Patent and Trademark Office, Technology Center 1600, Reception Area, 7th Floor, Examiner Name, Sequence Information, Crystal Mall One, 1911 South Clark Street, Arlington, VA 22202
   Or
   U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. **Federal Express, United Parcel Service, or other delivery service to:** U.S. Patent and Trademark Office, Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 01/29/2002

*O1PE*

## ERROR DETECTED          SUGGESTED CORRECTION          SERIAL NUMBER: 10/001,486

**ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE**

1____ Wrapped Nucleics
Wrapped Aminos

The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."

2____ Invalid Line Length

The rules require that a line not exceed 72 characters in length. This includes white spaces.

3 ✓ Misaligned Amino
Numbering

The numbering under each 5$^{th}$ amino acid is misaligned. Do not use tab codes between numbers; use **space characters**, instead.

4____ Non-ASCII

The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.

5____ Variable Length

Sequence(s)____ contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.

6____ PatentIn 2.0
"bug"

A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. **This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.**

7____ Skipped Sequences
(OLD RULES)

Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence:
(2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
(i)          SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
This sequence is intentionally skipped

Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.

8____ Skipped Sequences
(NEW RULES)

Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence.
<210> sequence id number
<400> sequence id number
000

9____ Use of n's or Xaa's
(NEW RULES)

Use of n's and/or Xaa's have been detected in the Sequence Listing.
Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.
In <220> to <223> section, please explain location of n or Xaa; and which residue n or Xaa represents.

10____ Invalid <213>
Response

Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence

11____ Use of <220>

Sequence(s) _____ missing the <220> "Feature" and associated numeric identifiers and responses. Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.
(See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)

12____ PatentIn 2.0
"bug"

Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.

13____ Misuse of n

n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.

AMC/MH -- Biotechnology Systems Branch -- 08/21/2001

OIPE

RAW SEQUENCE LISTING                          DATE: 07/08/2002
PATENT APPLICATION:  US/10/011,486A          TIME: 10:17:42

Input Set : N:\Crf3\Refhold\J001486.raw
Output Set: N:\CRF3\07082002\J011486A.raw

**SEQUENCE LISTING**
```
 1 (1) GENERAL INFORMATION:
 2      (i) APPLICANT: The Government of the United States of America
 3                     as respresented by the Secretary
 4                     Department of Health and Human Services
 5                     Washington, D.C.
 6                     Htun Ph.D., Han
 7                     Hager Ph.D., Gordon L.
 8     (ii) TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MONITORING
 9                              DNA BINDING MOLECULES IN LIVING CELLS
10    (iii) NUMBER OF SEQUENCES: 9
11     (iv) CORRESPONDENCE ADDRESS:
12          (A) ADDRESSEE: Needle & Rosenberg
13          (B) STREET: 127 Peachtree Street, Suite 1200
14          (C) CITY: Atlanta
15          (D) STATE: Georgia
16          (E) COUNTRY: USA
17          (F) ZIP: 30303
18      (v) COMPUTER READABLE FORM:
19          (A) MEDIUM TYPE: Floppy disk
20          (B) COMPUTER: IBM PC compatible
21          (C) OPERATING SYSTEM: PC-DOS/MS-DOS
22          (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
23     (vi) CURRENT APPLICATION DATA:
C--> 24         (A) APPLICATION NUMBER: US/10/011,486A
C--> 25         (B) FILING DATE: 30-Apr-2002
26          (C) CLASSIFICATION:
27    (vii) PRIOR APPLICATION DATA:
28          (A) APPLICATION NUMBER: 60/008,373
29          (B) FILING DATE: 08 Dec 1995
30   (viii) ATTORNEY/AGENT INFORMATION:
31          (A) NAME: Selby, Elizabeth
32          (B) REGISTRATION NUMBER: 38298
33          (C) REFERENCE/DOCKET NUMBER: 14014.0183
34     (ix) TELECOMMUNICATION INFORMATION:
35          (A) TELEPHONE: 404-688-0770
36          (B) TELEFAX: 404-688-9880
```

ERRORED SEQUENCES

```
320 (2) INFORMATION FOR SEQ ID NO: 2:
321      (i) SEQUENCE CHARACTERISTICS:
```

RAW SEQUENCE LISTING                          DATE: 07/08/2002
PATENT APPLICATION:  US/10/011,486A          TIME: 10:17:42

Input Set : N:\Crf3\Refhold\J001486.raw
Output Set: N:\CRF3\07082002\J011486A.raw

```
322          (A) LENGTH: 1071 amino acids
323          (B) TYPE: amino acid
324          (D) TOPOLOGY: linear
325    (ii) MOLECULE TYPE: protein
326    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
327    Met Ala His His His His His His Gly Tyr Pro Tyr Asp Val Pro Asp
328    1                   5                   10                  15
329    Tyr Ala Gln Ser Ser Ala Met Ser Lys Gly Glu Glu Leu Phe Thr Gly
330                    20                  25                  30
331    Val Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly His Lys
332                    35                  40                  45
333    Phe Ser Val Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu
334                    50                  55                  60
335    Thr Leu Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro
336    65                  70                  75                  80
337    Thr Leu Val Thr Thr Phe Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr
338                    85                  90                  95
339    Pro Asp His Met Lys Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu
340                    100                 105                 110
341    Gly Tyr Val Gln Glu Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr
342                    115                 120                 125
343    Lys Thr Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg
344                    130                 135                 140
345    Ile Glu Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly
346    145                 150                 155                 160
347    His Lys Leu Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala
348                    165                 170                 175
349    Asp Lys Gln Lys Asn Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn
350                    180                 185                 190
351    Ile Glu Asp Gly Ser Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr
352                    195                 200                 205
353    Pro Ile Gly Asp Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser
354    210                 215                 220
355    Thr Gln Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met
356    225                 230                 235                 240
357    Val Leu Leu Glu Phe Val Thr Ala Ala Gly Ile Thr His Gly Met Asp
358                    245                 250                 255
359    Glu Leu Tyr Lys Gly Ala Gly Ala Gly Ala Gly Ala Gly Ala Ile Ser
360                    260                 265                 270
361    Ala Leu Ile Leu Asp Ser Lys Glu Ser Leu Ala Pro Pro Gly Arg Asp
362                    275                 280                 285
363    Glu Val Pro Gly Ser Leu Leu Gly Gln Gly Arg Gly Ser Val Met Asp
364                    290                 295                 300
365    Phe Tyr Lys Ser Leu Arg Gly Gly Ala Thr Val Lys Val Ser Ala Ser
366    305                 310                 315                 320
367    Ser Pro Ser Val Ala Ala Ala Ser Gln Ala Asp Ser Lys Gln Gln Arg
368                    325                 330                 335
369    Ile Leu Leu Asp Phe Ser Lys Gly Ser Thr Ser Asn Val Gln Gln Arg
370                    340                 345                 350
```

```
       371   Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln
       372           355                 360                 365
       373   Gln Gln Gln Pro Gly Leu Ser Lys Ala Val Ser Leu Ser Met Gly Leu
       374           370                 375                 380
       375   Tyr Met Gly Glu Thr Glu Thr Lys Val Met Gly Asn Asp Leu Gly Tyr
       376   385                     390                 395                 400
       377   Pro Gln Gln Gly Gln Leu Gly Leu Ser Ser Gly Glu Thr Asp Phe Arg
       378           405                 410                 415
       379   Leu Leu Glu Glu Ser Ile Ala Asn Leu Asn Arg Ser Thr Ser Val Pro
       380           420                 425                 430
       381   Glu Asn Pro Lys Ser Ser Thr Ser Ala Thr Gly Cys Ala Thr Pro Thr
       382           435                 440                 445
       383   Glu Lys Glu Phe Pro Lys Thr His Ser Asp Ala Ser Ser Glu Gln Gln
       384           450                 455                 460
       385   Asn Arg Lys Ser Gln Thr Gly Thr Asn Gly Gly Ser Val Lys Leu Tyr
       386   465                     470                 475                 480
       387   Pro Thr Asp Gln Ser Thr Phe Asp Leu Leu Lys Asp Leu Glu Phe Ser
E--> 388               485                 490                 495
       389   Ala Gly Ser Pro Ser Lys Asp Thr Asn Glu Ser Pro Trp Arg Ser Asp
E--> 390               500                 505                 510
       391   Leu Leu Ile Asp Glu Asn Leu Leu Ser Pro Leu Ala Gly Glu Asp Asp
E--> 392           515                 520                 525
       393   Pro Phe Leu Leu Glu Gly Asn Thr Asn Glu Asp Cys Lys Pro Leu Ile
E--> 394           530                 535                 540
       395   Leu Pro Asp Thr Lys Pro Lys Ile Lys Asp Thr Gly Asp Thr Ile Leu
E--> 396   545                     550                 555                 560
       397   Ser Ser Pro Ser Ser Val Ala Leu Pro Gln Val Lys Thr Glu Lys Asp
E--> 398               565                 570                 575
       399   Asp Phe Ile Glu Leu Cys Thr Pro Gly Val Ile Lys Gln Glu Lys Leu
E--> 400               580                 585                 590
       401   Gly Pro Val Tyr Cys Gln Ala Ser Phe Ser Gly Thr Asn Ile Ile Gly
E--> 402           595                 600                 605
       403   Asn Lys Met Ser Ala Ile Ser Val His Gly Val Ser Thr Ser Gly Gly
E--> 404           610                 615                 620
       405   Gln Met Tyr His Tyr Asp Met Asn Thr Ala Ser Leu Ser Gln Gln Gln
E--> 406   625                     630                 635                 640
       407   Asp Gln Lys Pro Val Phe Asn Val Ile Pro Pro Ile Pro Val Gly Ser
E--> 408               645                 650                 655
       409   Glu Asn Trp Asn Arg Cys Gln Gly Ser Gly Glu Asp Ser Leu Thr Ser
E--> 410               660                 665                 670
       411   Leu Gly Ala Leu Asn Phe Pro Gly Arg Ser Val Phe Ser Asn Gly Tyr
E--> 412           675                 680                 685
       413   Ser Ser Pro Gly Met Arg Pro Asp Val Ser Ser Pro Pro Ser Ser Ser
E--> 414           690                 695                 700
       415   Ser Ala Ala Thr Gly Pro Pro Pro Lys Leu Cys Leu Val Cys Ser Asp
E--> 416   705                     710                 715                 720
       417   Glu Ala Ser Gly Cys His Tyr Gly Val Leu Thr Cys Gly Ser Cys Lys
E--> 418               725                 730                 735
       419   Val Phe Phe Lys Arg Ala Val Glu Gly Gln His Asn Tyr Leu Cys Ala
```

RAW SEQUENCE LISTING                    DATE: 07/08/2002
PATENT APPLICATION:   US/10/011,486A     TIME: 10:17:42

Input Set : N:\Crf3\Refhold\J001486.raw
Output Set: N:\CRF3\07082002\J011486A.raw

```
E--> 420                    740              745                750
     421       Gly Arg Asn Asp Cys Ile Ile Asp Lys Ile Arg Arg Lys Asn Cys Pro
E--> 422                755              760                765
     423       Ala Cys Arg Tyr Arg Lys Cys Leu Gln Ala Gly Met Asn Leu Glu Ala
E--> 424            770              775                780
     425       Arg Lys Thr Lys Lys Lys Ile Lys Gly Ile Gln Gln Ala Thr Ala Gly
E--> 426       785              790                795                800
     427       Val Ser Gln Asp Thr Ser Glu Asn Pro Asn Lys Thr Ile Val Pro Ala
E--> 428                805              810                815
     429       Ala Leu Pro Gln Leu Thr Pro Thr Leu Val Ser Leu Leu Glu Val Ile
E--> 430                820              825                830
     431       Glu Pro Glu Val Leu Tyr Ala Gly Tyr Asp Ser Ser Val Pro Asp Ser
E--> 432                835              840                845
     433       Ala Trp Arg Ile Met Thr Thr Leu Asn Met Leu Gly Gly Arg Gln Val
E--> 434            850              855                860
     435       Ile Ala Ala Val Lys Trp Ala Lys Ala Ile Leu Gly Leu Arg Asn Leu
E--> 436       865              870                875                880
     437       His Leu Asp Asp Gln Met Thr Leu Leu Gln Tyr Ser Trp Met Phe Leu
E--> 438                885              890                895
     439       Met Ala Phe Ala Leu Gly Trp Arg Ser Tyr Arg Gln Ser Ser Gly Asn
E--> 440                900              905                910
     441       Leu Leu Cys Phe Ala Pro Asp Leu Ile Ile Asn Glu Gln Arg Met Ser
E--> 442                915              920                925
     443       Leu Pro Gly Met Tyr Asp Gln Cys Lys His Met Leu Phe Val Ser Ser
E--> 444            930              935                940
     445       Glu Leu Gln Arg Leu Gln Val Ser Tyr Glu Glu Tyr Leu Cys Met Lys
E--> 446       945              950                955                960
     447       Thr Leu Leu Leu Leu Ser Ser Val Pro Lys Glu Gly Leu Lys Ser Gln
E--> 448                965              970                975
     449       Glu Leu Phe Asp Glu Ile Arg Met Thr Tyr Ile Lys Glu Leu Gly Lys
E--> 450                980              985                990
     451       Ala Ile Val Lys Arg Glu Gly Asn Ser Ser Gln Asn Trp Gln Arg Phe
E--> 452                995              1000               1005
     453       Tyr Gln Leu Thr Lys Leu Leu Asp Ser Met His Glu Val Val Glu Asn
E--> 454            1010             1015               1020
     455       Leu Leu Thr Tyr Cys Phe Gln Thr Phe Leu Asp Lys Thr Met Ser Ile
E--> 456       1025             1030               1035               1040
     457       Glu Phe Pro Glu Met Leu Ala Glu Ile Ile Thr Asn Gln Ile Pro Lys
E--> 458                1045             1050               1055
E--> 459       Tyr Ser Asn Gly Asn Ile Lys Lys Leu Leu Phe His Gln Lys   *
E--> 460                1060             1065               1070
```

**VERIFICATION SUMMARY**                    DATE: 07/08/2002
PATENT APPLICATION:   US/10/011,486A        TIME: 10:17:43

Input Set : **N:\Crf3\Refhold\J001486.raw**
Output Set: **N:\CRF3\07082002\J011486A.raw**

L:24 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:25 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:43 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=1
L:388 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:2
M:332 Repeated in SeqNo=2
L:459 M:342 E: Invalid Stop Code On Error, STOP CODON:*
L:468 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=3
L:478 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=4
L:488 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=5
L:498 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=6
L:508 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=7
L:518 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=8